## **RAW SEQUENCE LISTING**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: Date Processed by STIC:

## ENTERED



PCT

RAW SEQUENCE LISTING DATE: 12/09/2004
PATENT APPLICATION: US/10/516,476 TIME: 14:25:28

```
Input Set : A:\Avalon 165.txt
                     Output Set: N:\CRF4\12092004\J516476.raw
      4 <110> APPLICANT: Avalon Pharmaceuticals
      6 <120> TITLE OF INVENTION: Cancer-Linked Gene as Target for Chemotherapy
      8 <130> FILE REFERENCE: 689290-165
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/516,476
C--> 11 <141> CURRENT FILING DATE: 2004-11-30
     13 <150> PRIOR APPLICATION NUMBER: US/60/385,505
     14 <151> PRIOR FILING DATE: 2002-06-04
     16 <160> NUMBER OF SEQ ID NOS: 39
     18 <170> SOFTWARE: PatentIn version 3.0
     20 <210> SEO ID NO: 1
     21 <211> LENGTH: 4567
     22 <212> TYPE: DNA
     23 <213> ORGANISM: Homo sapiens
     25 <400> SEQUENCE: 1
     26 geggeegeee catteecaga eeggeegeea geecatetgg ttageteeeg eegeteegeg
                                                                               60
     27 cegeceggga gtegggagee geggggaace gggeacetge accegeetet gggagtgagt
                                                                              120
     28 ggttccagct ggtgcctggc ctgtgtctct tggatgccct gtggcttcag tccgtctcct
                                                                              180
     29 gttgcccacc acctcgtccc tgggccgcct gataccccag cccaacagct aaggtgtgga
                                                                              240
     30 tggacagtag ggggctggct teteteactg gteaggggte tteteceetg tetgeeteee
     31 ggagetagga etgeagaggg geetateatg gtgettgeag geeceetgge tgtetegetg
     32 ttgctgccca gcctcacact gctggtgtcc cacctctcca gctcccagga tgtctccagt
     33 gageceagea gtgageagea getgtgegee ettageaage acceeacegt ggeetttgaa
     34 gacctgcage egtgggtete taacttcace taccetggag eeegggattt eteccagetg
```

300 360 420 480 540 35 getttggace ceteegggaa eeageteate gtgggageea ggaactaeet etteagaete 600 36 ageettgeea atgtetetet tetteaggee acagagtggg cetecagtga ggacaegege 660 720 37 cgctcctgcc aaagcaaagg gaagactgag gaggagtgtc agaactacgt gcgagtcctg 780 38 atogtogoog googgaaggt gttoatgtgt ggaaccaatg cottttoccc catgtgcacc 39 agcagacagg tggggaacct cagccggact actgagaaga tcaatggtgt ggcccgctgc 840 40 ccctatgacc cacqccacaa ctccacagct gtcatctcct cccaggggga gctctatgca 900 41 gccacggtca tcgacttctc aggtcgggac cctgccatct accgcagcct gggcagtggg 960 42 ccaccgcttc gcactgccca atataactcc aagtggctta atgagccaaa cttcgtggca 1020 1080 43 gcctatgata ttgggctgtt tgcatacttc ttcctgcggg agaacgcagt ggagcacgac 44 tgtggacgca ccgtgtactc tcgcgtggcc cgcgtgtgca agaatgacgt ggggggccga 1140 1200 45 ttcctgctgg aggacacatg gaccacattc atgaaggccc ggctcaactg ctcccgcccg 1260 46 ggcgaggtcc ccttctacta taacgagctg cagagtgcct tccacttgcc agagcaggac 47 ctcatctatg gagttttcac aaccaacgta aacagcatcg cggcttctgc tgtctgcgcc 1320 48 ttcaacctca gtgctatctc ccaggctttc aatggcccat ttcgctacca ggagaacccc 1380 49 agggctgcct ggctccccat agccaacccc atccccaatt tccagtgtgg caccctgcct 1440 1500 50 gagaceggte ceaacgagaa cetgaeggag egeageetge aggaegegea gegeetette 1560 51 ctgatgagcg aggccgtgca gccggtgaca cccgagccct gtgtcaccca ggacagcgtg 1620 52 cgcttctcac acctcgtggt ggacctggtg caggctaaag acacgctcta ccatgtactc 1680 53 tacattggca ccgagtcggg caccatcctg aaggcgctgt ccacggcgag ccgcagcctc

54 cacggetget acctggagga getgeacgtg etgeceeeg ggegeegega geeeetgege

1740





## RAW SEQUENCE LISTING DATE: 12/09/2004 PATENT APPLICATION: US/10/516,476 TIME: 14:25:28

Input Set : A:\Avalon 165.txt

Output Set: N:\CRF4\12092004\J516476.raw

55	agcctgcgca	tcctgcacag	cgcccgcgcg	ctcttcgtgg	ggctgagaga	cggcgtcctg	1800		
56	cgggtcccac	tggagaggtg	cgccgcctac	cgcagccagg	gggcatgcct	gggggcccgg	1860		
57	gacccgtact	gtggctggga	cgggaagcag	caacgttgca	gcacactcga	ggacagctcc	1920		
58	aacatgagcc	tctggaccca	gaacatcacc	gcctgtcctg	tgcggaatgt	gacacgggat	1980		
59	gggggcttcg	gcccatggtc	accatggcaa	ccatgtgagc	acttggatgg	ggacaactca	2040		
60	ggctcttgcc	tgtgtcgagc	tcgatcctgt	gattcccctc	gaccccgctg	tgggggcctt	2100		
61	gactgcctgg	ggccagccat	ccacatcgcc	aactgctcca	ggaatggggc	gtggaccccg	2160		
62	tggtcatcgt	gggcgctgtg	cagcacgtcc	tgtggcatcg	gcttccaggt	ccgccagcga	2220		
63	agttgcagca	accctgctcc	ccgccacggg	ggccgcatct	gcgtgggcaa	gagccgggag	2280		
64	gaacggttct	gtaatgagaa	cacgccttgc	ccggtgccca	tcttctgggc	ttcctggggc	2340		
65	tcctggagca	agtgcagcag	caactgtgga	gggggcatgc	agtcgcggcg	tcgggcctgc	2400		
			gggctgcggc				2460		
			caccccctgg				2520		
68	ggcggggcac	ggcaggagca	gcggttccgc	ttcacctgcc	gcgcgcccct	tgcagacccg	2580		
			gagaaggacc				2640		
			cctggtggag				2700		
			ggccgcctgg				2760		
			caagagaacg				2820		
			tgccgagtac				2880		
			gacctcatgg				2940		
			ctgcaccagc				3000		
			ggcactatgt				3060		
			gtgcactgac				3120		
			cagcgcmtgt				3180		
			catcctgcca				3240		
			cttggtggcc				3300		
			cctgtcttgc				3360		
			ccccaaccat				3420		
			ggaattcaag				3480		
			attgcagcag				3540		
			cttccggccc				3600		
			tggggacttg				3660		
			ccagtttggt				3720		
			cccatccggc				3780		
			gctgacatat				3840		
			ctgtcacgct				3900		
			ytggagaatg				3960		
			tggccccac				4020		
			tcctgattat				4080		
			gggctgggcc				4140		
			ccacctctgc				4200		
			tactgggaac				4260		
			tcagccttgc				4320		
			cagaagtctg				4380		
			gttggctgag				4440		
						a caatttgcta	4500		
						ccagtcaaaa	4560		
	aaaaact		, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		- auguegueut		4567		
105 <210> SEQ ID NO: 2									
TO2 <2TO> 2DQ TD NO: 7									





## RAW SEQUENCE LISTING DATE: 12/09/2004 PATENT APPLICATION: US/10/516,476 TIME: 14:25:28

Input Set : A:\Avalon 165.txt

Output Set: N:\CRF4\12092004\J516476.raw

106	<211> LENGTH: 453										
	<212> TYPE: DNA										
	<213> ORGAN		sapiens								
110	<400> SEQUENCE: 2										
111	agtaatcagc	tcggtaccgg	catgtgctgt	agccagcgca	ggttatccac	ctctgcctgg	60				
112	aagtctatac	tctacccagg	gcatccctct	ggtcagaggc	agtgagtact	gggaactgga	120				
113	ggctgacctg	tgcttagaag	tcctttaatc	tgggctggta	caggcctcag	ccttgccctc	180				
114	aatgcacgaa	aggtggccca	ggagagagga	tcaatgccat	aggaggcaga	agtctggcct	240				
115	ctgtgcctct	atggagacta	tcttccagtt	gctgctcaac	agagttgttg	gctgagacct	300				
116	gcttgggagt	ctctgctggc	ccttcatctg	ttcaggaaca	cacacacaca	cacactcaca	360				
117	cacgcacaca	caatcacaat	ttgctacagc	aacaaaaaag	acattgggct	gtggcattat	420				
			tcaaaaaaaa				453				
	<210> SEQ ID NO: 3										
122	<211> LENGT	TH: 4675									
123	<212> TYPE:	DNA									
124	<213> ORGAN	NISM: Homo s	sapiens								
	<400> SEQUE		-								
			ccggccgcca	gcccatctgg	ttagctcccg	ccgctccgcg	60				
			gcggggaacc				120				
			ctgtgtctct				180				
			tgggccgcct				240				
			tctctcactg				300				
			gggacctggg				360				
			tggaggtctt				420				
			gcttgcaggc				480				
			cctctccagc				540				
			tagcaagcac				600				
			ccctggagcc				660				
			gggagccagg				720				
			agagtgggcc				780				
			ggagtgtcag				840				
			aaccaatgcc				900				
			tgagaagatc				960				
			catctcctcc				1020				
			tgccatctac				1080				
			gtggcttaat				1140				
			cctgcgggag				1200				
			cgtgtgcaag				1260				
			gaaggcccgg				1320				
			gagtgccttc				1380				
			cagcatcgcg				1440				
			tggcccattt				1500				
			ccccaatttc				1560				
			cagcctgcag				1620				
							1680				
			cgagccctgt				1740				
			ggctaaagac				1800				
			ggcgctgtcc				1860				
			gcccccggg								
128	ccgcacageg	cccgcgcgct	cttcgtgggg	ctgagagacg	gegreergeg	ggreecactg	1920				





RAW SEQUENCE LISTING DATE: 12/09/2004
PATENT APPLICATION: US/10/516,476 TIME: 14:25:28

Input Set : A:\Avalon 165.txt

Output Set: N:\CRF4\12092004\J516476.raw

```
159 gagaggtgcg ccgcctaccg cagccagggg gcatgcctgg gggcccggga cccgtactgt
                                                                         1980
                                                                         2040
160 ggctgggacg ggaagcagca acgttgcagc acactcgagg acagctccaa catgagcctc
161 tggacccaga acatcaccgc ctgtcctgtg cggaatgtga cacgggatgg gggcttcggc
                                                                         2100
162 ccatqqteac catqqcaacc atqtqaqcac ttqqatqqqq acaactcaqq ctcttqcctq
                                                                         2160
163 tgtcgagete gateetgtga tteecetega eeeegetgtg ggggeettga etgeetgggg
                                                                         2220
164 ccagccatcc acategecaa etgetecagg aatggggegt ggacccegtg gtcategtgg
                                                                         2280
165 gegetgtgea geaegteetg tggeategge tteeaggtee geeagegaag ttgeageaae
                                                                         2340
166 cctgctcccc gccacggggg ccgcatctgc gtgggcaaga gccgggagga acggttctgt
                                                                         2400
167 aatgagaaca cgccttgccc ggtgcccatc ttctgggctt cctggggctc ctggagcaag
                                                                         2460
                                                                         2520
168 tqcaqcaqca actqtqqaqq qqqcatqcaq tcqcqgcqtc qgqcctqcqa gaacgqcaac
169 tcctgcctgg gctgcggcgt ggagttcaag acgtgcaacc ccgagggctg ccccgaagtg
                                                                         2580
170 cggcgcaaca ccccctggac gccgtggctg cccgtgaacg tgacgcaggg cggggcacgg
                                                                         2640
171 caggageage ggtteegett cacetgeege gegeeeettg cagaeeegea eggeetgeag
                                                                         2700
172 ttcggcagga gaaggaccga gacgaggacc tgtcccgcgg acggctccgg ctcctgcgac
                                                                         2760
173 accgaegece tggtggaggt cetectgege agegggagea ceteceegea caeggtgage
                                                                         2820
174 gggggctggg ccgcctgggg cccgttggtcg tcctgctccc gggactgcga gctgggcttc
                                                                         2880
                                                                         2940
175 cgcgtccgca agagaacgtg cactaacccg gagccccgca acgggggcct gccctgcgtg
176 ggcgatgctg ccgagtacca ggactgcaac ccccaggctt gcccagttcg gggtgcttgg
                                                                         3000
177 teetgetgga ceteatggte teeatgetea getteetgtg gtgggggtea etateaaege
                                                                         3060
178 acception geaccageee egeaccetee ecaggigagg acateigtei egggeigeae
                                                                         3120
                                                                         3180
179 acggaggagg cactatgtgc cacacaggcc tgcccagaag gctggtcgcc ctggtctgag
180 tggagtaagt gcactgacga cggagcccag agccgaagcc ggcactgtga ggagctcctc
                                                                         3240
181 ccagggtcca gcgcmtgtgc tggaaacagc agccagagcc gcccctgccc ctacagcgag
                                                                         3300
182 attesegtea teetgeeage etceageatg gaggaggeea cegaetgtge agggtteaat
                                                                         3360
                                                                         3420
183 ctcatccact tgqtqqccac qqqcatctcc tqcttcttqq qctctqqqct cctgacccta
184 gcagtgtacc tgtcttgcca gcactgccag cgtcagtccc aggagtccac actggtccat
                                                                         3480
185 cctgccaccc ccaaccattt gcactacaag ggcggaggca ccccgaagaa tgaaaagtac
                                                                         3540
                                                                         3600
186 acacccatgg aattcaagac cctgaacaag aataacttga tccctgatga cagagccaac
187 ttctacccat tgcagcagac caatgtgtac acgactactt actacccaag ccccctgaac
                                                                         3660
                                                                         3720
188 aaacacaget teeggeeega ggeeteaeet ggacaaeggt getteeecaa eagetgatae
189 cgccgtcctg gggacttggg cttcttgcct tcataaggca cagagcagat ggagatggga
                                                                         3780
190 cagtggagec agtttggttt tetecetetg caetaggeca agaacttget geettgeetg
                                                                         3840
191 tggggggtcc catccggctt cagagagctc tggctggcat tgaccatggg ggaaagggct
                                                                         3900
192 ggtttcagge tgacatatgg eegeaggtee agttcageee aggtetmtea tggttatett
                                                                         3960
193 ccaacccact gtcacgctga cactatgctg ccatgcctgg gctgtggacc tactgggcat
                                                                         4020
                                                                         4080
194 ttgaggaayt ggagaatgga gatggcaaga gggcaggctt ttaagtttgg gttggagaca
195 acttectgtg gececeacaa getgagtetg geetteteea getggeeeca aaaaaggeet
                                                                         4140
196 ttgctacatc ctgattatct ctgaaagtaa tcaatcaagt ggctccagta gctctggatt
                                                                         4200
                                                                         4260
197 ttctgccagg gctgggccat tgtggtgctg ccccagtatg acatgggacc aaggccagcg
198 caggttatcc acctctgcct ggaagtctat actctaccca gggcatccct ctggtcagag
                                                                         4320
199 gcagtgagta ctgggaactg gaggctgacc tgtgcttaga agtcctttaa tctgggctgg
                                                                         4380
200 tacaggeete ageettgeee teaatgeacg aaaggtggee caggagagag gateaatgee
                                                                         4440
201 ataggaggca gaagtetgge etetgtgeet etatggagae tatetteeag ttgetgetea
                                                                         4500
202 acagagttqt tqqctqaqac ctqcttggqa qtctctqctg qcccttcatc tgttcaggaa
                                                                         4560
203 cacacacaca cacacactca cacacgcaca cacaatcaca atttgctaca gcaacaaaaa
                                                                         4620
204 agacattggg ctgtggcatt attaattaaa gatgatatcc agtcaaaaaa aaact
                                                                         4675
```

207 <210> SEQ ID NO: 4 208 <211> LENGTH: 4731

209 <212> TYPE: DNA





RAW SEQUENCE LISTING DATE: 12/09/2004
PATENT APPLICATION: US/10/516,476 TIME: 14:25:28

Input Set : A:\Avalon 165.txt

Output Set: N:\CRF4\12092004\J516476.raw

210 <213 > ORGANISM: Homo sapiens 212 <400> SEQUENCE: 4 213 attggagatg ctcgggggca ggctgccgcg ttgtgtcctg cttttctgcg gccagaccaa 60 214 qccqtctqqa qctqctqqtc aqqttttctt gctgacctca cctgaccaca gtggcctggg 120 215 tggactctac agggaaatgt tgttttctcc ctgggagcag tagcagcagt cctggctccc 240 216 ctqqactqaq aactcctcat caqccccaqq aaqcccggac cccctttcag ggatctggaa 217 cogqtqtqcc tqtqqcccca qqtctqctcc caqqcqtqqq ctqaaqtcct gacttctgtc 300 218 gctgggggca aggagtggga gagcccagct gctgcctggg ctttggcaga cagcaggctg 360 219 atggtgctgg cttccccgag actgcttctc ctgcctgctg tctgatttcc ctgcatggtg 420 220 cccqcaqctq agctqctacq ggtcttctcc cctgtctgcc tcccggagct aggactgcag 480 221 aggggeetat catggtgett geaggeecee tggetgtete getgttgetg eecageetea 540 222 cactgctggt gtcccacctc tccagctccc aggatgtctc cagtgagccc agcagtgagc 600 223 aqcaqctqtq cqcccttagc aaqcacccca ccgtggcctt tgaagacctg cagccgtggg 660 224 tetetaaett eacetaeeet qqaqeeeggg attteteeea getggetttg gaeeeeteeg 720 225 ggaaccagct catcgtggga gccaggaact acctcttcag actcagcctt gccaatgtct 780 226 ctcttcttca ggccacagag tgggcctcca gtgaggacac gcgccgctcc tgccaaagca 840 227 aagggaagac tgaggaggag tgtcagaact acgtgcgagt cctgatcgtc gccggccgga 900 228 aggtgttcat gtgtggaacc aatgcctttt cccccatgtg caccagcaga caggtgggga 960 229 acctcagccg gactactgag aagatcaatg gtgtggcccg ctgcccctat gacccacgcc 1020 230 acaactccac agctgtcatc tcctcccagg gggagctcta tgcagccacg gtcatcgact 1080 231 teteaggteg ggaceetgee atetacegea geetgggeag tgggceaceg ettegeactg 1140 232 cccaatataa ctccaagtgg cttaatgagc caaacttcgt ggcagcctat gatattgggc 1200 233 tgtttgcata cttcttcctg cgggagaacg cagtggagca cgactgtgga cgcaccgtgt 1260 234 actetegeqt qqcccqcqtq tqcaaqaatq acqtgggggg ccgatteetg ctggaggaca 1320 235 catqqaccac attcatqaaq qcccqqctca actqctcccq cccgggcgag gtccccttct 1380 236 actataacga gctgcagagt gccttccact tgccagagca ggacctcatc tatggagttt 1440 237 tcacaaccaa cgtaaacagc atcgcggctt ctgctgtctg cgccttcaac ctcagtgcta 1500 238 teteccaqqe tttcaatqqc ccattteqct accaqqaqaa ccccagggct gcctggctcc 1560 1620 239 ccatagecaa ceccatecee aatttecagt gtggcaceet geetgagace ggteccaaeg 240 agaacctgac ggagcgcagc ctgcaggacg cgcagcgcct cttcctgatg agcgaggccg 1680 241 tgcagccggt gacacccgag ccctgtgtca cccaggacag cgtgcgcttc tcacacctcg 1740 242 tggtggacct ggtgcaggct aaagacacgc tctaccatgt actctacatt ggcaccgagt 1800 1860 243 egggcaccat cetgaaggeg etgtecaegg egageegeag cetecaegge tgetacetgg 244 aggagetgea egtgetgeec eeegggegee gegageeeet gegeageetg egeateetge 245 acagegeeg egegetette gtggggetga gagaeggegt eetgegggte eeactggaga 1980 246 qqtqcqccqc ctaccqcaqc cagggggcat gcctgggggc ccgggacccg tactgtggct 2040 247 gggacgggaa gcagcaacgt tgcagcacac tcgaggacag ctccaacatg agcctctgga 2100 248 cccagaacat caccgcctgt cctgtgcgga atgtgacacg ggatgggggc ttcggcccat 2160 249 ggtcaccatg gcaaccatgt gagcacttgg atggggacaa ctcaggctct tgcctgtgtc 2220 2280 250 qaqctcqatc ctgtqattcc cctcgacccc gctgtggggg ccttgactgc ctggggccag 251 ccatccacat cgccaactgc tccaggaatg gggcgtggac cccgtggtca tcgtgggcgc 2340 252 tqtqcaqcac qtcctqtqqc atcggcttcc aggtccgcca gcgaagttgc agcaaccctg 2400 2460 253 ctccccqcca cqqqqqccqc atctgcgtgg gcaagagccg ggaggaacgg ttctgtaatg 254 aqaacacqcc ttqcccqqtq cccatcttct gggcttcctg gggctcctgg agcaagtgca 2520 2580 255 gcagcaactg tggaggggc atgcagtcgc ggcgtcgggc ctgcgagaac ggcaactcct 2640 256 gcctgggctg cggcgtggag ttcaagacgt gcaaccccga gggctgcccc gaagtgcggc 2700 257 gcaacacccc ctggacgccg tggctgcccg tgaacgtgac gcagggcggg gcacggcagg 258 ageageggtt cegetteace tgeegegege ceettgeaga ceegeaegge etgeagtteg 2760 259 gcaggagaag gaccgagacg aggacctgtc ccgcggacgg ctccggctcc tgcgacaccg 2820





DATE: 12/09/2004

TIME: 14:25:29

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/516,476

Input Set : A:\Avalon 165.txt

Output Set: N:\CRF4\12092004\J516476.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date